Varying Portfolio Construction of Stocks Using Genetic Network Programming with Control Nodes

Etsushi Ohkawa
etsushi@fuji.waseda.jp

Yan Chen
chenyan15153
@akane.waseda.jp

Shingo Mabu
mabu@aoni.waseda.jp

Kaoru Shimada
k.shimada@aoni.waseda.jp

Kotaro Hirasawa
hirasawa@waseda.jp

Graduate School of Information, Production and Systems, Waseda University,
2-7 Hibikino, Wakamatsu-ku, Kitakyushu, Fukuoka 808-0135, Japan

ABSTRACT

A new evolutionary method named "Genetic Network Programming with Control Nodes, GNPcn" has been proposed and applied to determine the timing of buying and selling stocks. GNPcn represents its solution as a directed graph structure which has some useful features inherently. For example, GNPcn has the implicit memory function which memorizes the past action sequences of agents and GNPcn can re-use nodes repeatedly in the network flow, so highly compact graph structures can be made. GNPcn can improve the strategy of buying and selling stocks of multi issues. Its effectiveness is confirmed by some simulations.

Categories and Subject Descriptors

J.1 [**Computer Applications**]: Administrative Data Processing

General Terms

Algorithms

Keywords

Genetic Network Programming, evolutionary computation, Candlestick Chart, buying and selling stocks.

1. GENETIC NETWORK PROGRAMMING WITH CONTROL NODES

1.1 Structure of GNPcn

GNP represents its solution as a directed graph structure which has some useful features inherently. GNP consists of the following nodes.

Judgment node: Judgment node returns a judgment result on the information gained from the environments, and determines the next node.

Processing node: Processing node works as a processing function, and the next node is uniquely determined by its single connection.

GNPcn starts from one of the *control nodes*, and the activated node is transferred to one of the control nodes after executing a certain number of processing nodes. Consequently, the performance of GNPcn improves because the increase of the number of control nodes contributes to searching the solution space widely and finding many smart solutions.

Fig. 1 shows the basic structure of GNPcn. Several control nodes are added to the graph structure. GNPcn uses one of the groups of control nodes for one stock when dealing with multi-issues in the stock market.

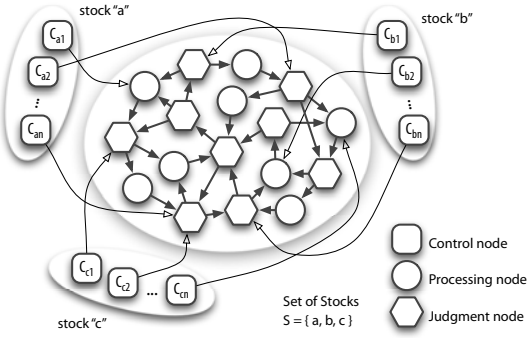

Figure 1: Structure of GNPcn

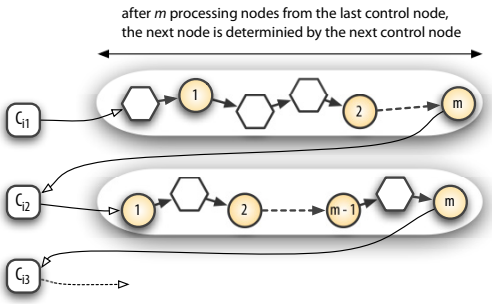

Figure 2: Transition of GNPcn

1.2 Node transition of GNPcn

For example, when GNPcn deals with stock "i", the node transition starts from control node "C_{i1}." The activated node returns to one of the control nodes after transiting m processing nodes from the last control node. That is, the activated node returns to C_{i2} firstly, then, to $C_{i3}, C_{i4}, \ldots, C_{in}, C_{i1}, C_{i2}, \ldots$ sequentially as shown in Fig. 2.

In "*one* day" period, the current activated node of each stock transits one node.

GECCO'08, July 12–16, 2008, Atlanta, Georgia, USA.
ACM 978-1-60558-130-9/08/07.

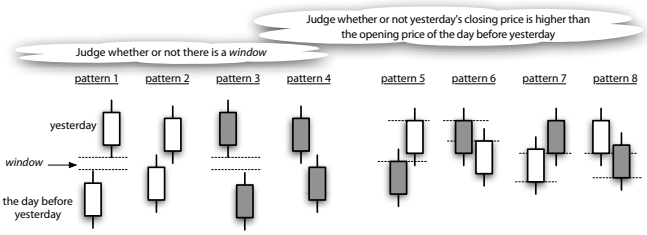

Figure 3: Judgment patterns

1.3 Evolution of GNPcn

The following procedure shows the evolution of GNPcn.

1. **Generate an initial population randomly.**
2. **Evaluation** — The stock trading for each issue is carried out and the fitness of each individual is calculated.
3. **Genetic Operation** — The elite individuals are preserved to next generation. And the rest is replaced by the new ones generated by crossover and mutation.
4. **Repeat 2, 3 until a certain generation.**

2. STRATEGY OF BUYING AND SELLING STOCKS OF MULTI ISSUES

2.1 Candlestick Chart

Candlestick Chart represents 4 stock prices — Opening, High, Low and Closing in a period. When the box is *black*, it means that the Closing price is lower than the Opening Price, and when *white*, it means the opposite.

2.2 Strategy of buying and selling stocks

In this paper, judgment nodes check the Candlestick Chart patterns (Fig. 3), and processing nodes work for buying or selling stocks. Therefore, a chain of judgment nodes and processing nodes expresses the strategy of buying/selling.

At the processing node, the trading is executed using the opening price of the day. When GNPcn determines to sell stock i, all of the stocks in hand are sold off. When GNPcn determines to buy stock i on day d, stock i is bought as much as possible using the money allocated to issue i, which is calculated by the distribution ratio $P(i,d)$ of the total fund in hand.

2.3 Distribution ratio to each stock

Distribution ratio $P(i,d)$, profit ratio $R(i,d)$ and capital gain $G(i,d)$ are calculated by the following equations:

$$P(i,d) = \frac{\exp\left(R(i,d) \,/\, T\right)}{\sum_{i \in S} \exp\left(R(i,d) \,/\, T\right)}$$

$$R(i,d) = \begin{cases} G(i,d) \,/\, Buy(i,d), & \text{if } Buy(i,d) \neq 0. \\ 0, & \text{if } Buy(i,d) = 0. \end{cases}$$

$$G(i,d) = Sell(i,d) + Price(i,d) * Unit(i,d) - Buy(i,d)$$

where,

$Sell(i,d)$, $Buy(i,d)$ **:** sum of the money earned by selling/buying the stocks of the issue name $i \in S$ until day d.

$Price(i,d)$ **:** stock price of issue i on day d.

$Unit(i,d)$ **:** amount of stocks of issue i in hand on day d.

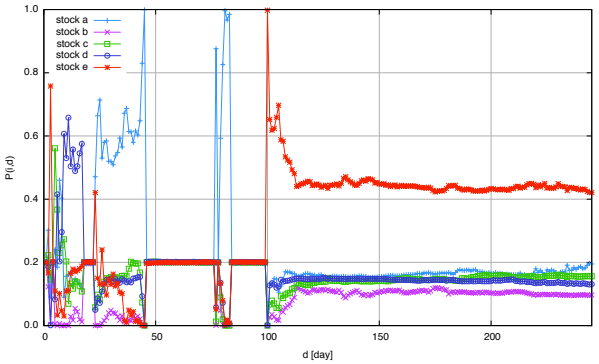

Figure 4: Distribution ratio $P(i,d)$ (Testing in 2003)

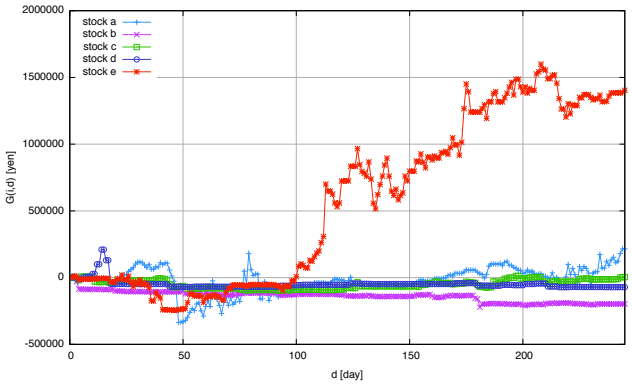

Figure 5: Capital gain $G(i,d)$ (Testing in 2003)

T **:** temperature parameter. S **:** set of suffixes of issues.

The purchase capital is evenly distributed to each stock initially, but the distribution ratios are updated every day.

In addition, during the evolution, GNPcn uses the sum of the capital gain $G(i,d_t)$ as a fitness value, where d_t indicates the final day of the dealing.

3. SIMULATIONS

In the simulations, the stock prices of the following period are used. — **Training:** 2002–2003, **Testing:** 2004.

The data used in the simulations are the stock data of 5 issues[1], all of which are listed in the First Section of the Tokyo Exchange market. The program used in the testing simulations is the best one at the last generation in the training. The initial fund is 5 million JPY. The capital gains achieved during the trading are also available for the next trading.

Fig. 4 shows the distribution ratio of the stock "e" is over 0.4 finally. It is found from Fig. 5 that the profit ratio obtained by GNPcn is mainly from the stock "e", while other brands do not gain profit so much.

The advantage of the proposed method is to determine the distribution ratio of the proper purchase capital to each stock automatically.

4. REFERENCES

[1] S. Mabu, K. Hirasawa and J. Hu, A Graph-Based Evolutionary Algorithm: Genetic Network Programming and Its Extension Using Reinforcement Learning, *Evolutionary Computation, MIT Press*, Vol. 15, No. 3, pp. 369–398, 2007.

[1] Meiji Seika, Ltd., Takeda Pharmaceutical Company Limited, Nippon Steel Corporation, NEC Corporation and Nissan Motor Co.,Ltd.